

63 Quantifying nationality bias in data from different social media platforms for visitor monitoring in Nikko National Park, Japan

Masahiro Kajikawa¹, Takafumi Miyasaka¹, Yutaka Kubota¹, Akihiro Oba², Katori Miyasaka³

¹Graduate School of Environmental Studies, Nagoya University, Japan. ²Research and Development, Initiative, Chuo University, Japan. ³College of Bioresource Sciences, Nihon University, Japan

Introduction

Nature-based tourism in protected areas has grown worldwide in recent years, but excessive use of natural areas can result in their degradation or loss. Visitor management should be based on proper monitoring data to achieve quality experiences for visitors without damaging nature resources.

Visitor data are typically collected through field surveys, but budget and human resource constraints can limit the spatio-temporal resolution of survey data. Geotagged photos and messages posted on social media by visitors have attracted attention as useful sources of information with high spatio-temporal resolution. Previous studies, however, have raised concerns that biases in social media data arising from the sociodemographic attributes of posters can create challenges in determining who and what the social media data represent, and in interpreting this data in a reliable way.

The present study focused on nationality bias, i.e., differences between the nationality of actual visitors versus the nationality of those visitors who post on social media. Nationality bias can arise due to significant differences in the use of social media from country to country (including the proportion of the population using social media). Cultural and values differences between countries can also greatly influence visitor behavior, and these gaps may lead to over- or under-estimation of visitors from specific countries or regions. However, nationality bias is not fully understood because few studies have explicitly considered it until now (e.g., Heikinheimo et al., 2017; Sinclair et al., 2020). In addition, the previous studies were confined to protected areas where the majority of visitors were from Europe.

This study examined one protected area in Japan to which many of the foreign visitors come from Asia. Our objectives were to quantify nationality bias in visitor-generated social media

data in order to improve general understanding of this issue, and to examine the feasibility of using social media for visitor monitoring and management in protected areas.

Methods

The study area was the Oku-Nikko area of Nikko National Park, one of Japan's most popular national parks. In October 2019, we conducted a questionnaire survey in four locations with different levels of popularity and different characteristics, to cover a diverse range of visitors to the area. Questions covered the nationality of visitors, and their usage of social media, including what social media platforms they used, whether they usually shared their travel experiences on social media, and whether they usually geotagged their posts. We also investigated other basic visitor attributes and behavior, including gender, age, size of party, means of transportation, length of stay, and places visited within the Oku-Nikko area.

When conducting surveys at multiple locations, sampling should be done according to the spatial distribution of the number of actual visitors at those locations, to avoid over- or under-estimating characteristics of some locations. Since such detailed reference visitor data as the number of visitors in different segments of a natural area are generally not available, we collected geotagged tweets posted around the four sampling locations in October 2019. We then estimated the relative ratio of actual visitor numbers at the locations using two parameters: the number of visitors who actually posted geotagged tweets (on Twitter) around each of the locations, and the percentage of respondents who answered they would post their travel experiences with geotags (on Twitter) at each location in our survey. According to the relative ratio, we performed a resampling of the dataset we collected in the field ($n = 1185$) and built a new dataset ($n = 552$).

To simplify analysis and interpretation of results, we considered social media posters to be only those visitors who used Twitter, Instagram, or Facebook platforms. For the same reasons, we classified nationalities into Japan, Asia (excluding Japan), and other countries (88.0% of which were composed of Europe and North America).

First, we compared visitor attributes (except nationality) and behaviors between different nationalities of visitors, and between all visitors and social media posters among them, in order to characterize these different groups. Then, we analyzed whether the ratio of visitors posting on social media to all visitors differed by nationality, to test for nationality bias in social media data.

Results & Discussion

We found that visitors posting on Instagram and Facebook and visitors from Asia and other countries generally had similar patterns of visitor attributes, except group size. The foreign visitors, particularly Asian visitors, also had similar patterns of visiting places compared to visitors posting on the two platforms. These attribute and behavior patterns were different from the patterns of all visitors, most of whom were Japanese (87.5%). For example, foreign visitors were younger, the group size of Asian visitors tended to be larger; the group size of visitors from other countries tended to be smaller, most

foreign visitors tended to be first-time visitors, and foreign visitors tended to visit high-accessibility places. In addition, the percentages of Instagram and Facebook posters were much higher for Asian visitors and lower for Japanese visitors, compared to all visitors. These results indicated that foreign visitors had different characteristics from all visitors or Japanese visitors and would be overestimated in Instagram and Facebook data. On the other hand, visitors posting on Twitter did not significantly differ in visitor attributes or behaviors from all visitors, except in terms of age. This could be explained by the fact that Twitter was a more popular platform for Japanese than Instagram and Facebook, which is the opposite of the international trend. The result suggests that in Japan Twitter would be more representative of all visitors in the areas where most of the visitors are domestic (Japanese).

Our findings would vary depending on the status of use of social media, which can differ from country to country. Further studies are needed to elucidate nationality bias in different countries and to improve the feasibility of using social media to understand visitor to protected areas.

References

Heikinheimo et al. 2017. <https://doi.org/10.3390/ijgi6030085>. Sinclair et al. 2020. <https://doi.org/10.1016/j.jenvman.2020.110418>.